Sorting Out a Promiscuous Superfamily: Towards Cadherin Connectomics

Sorting Out a Promiscuous Superfamily:

Towards Cadherin Connectomics

Marcos Sotomayor[1], Rachelle Gaudet[2], and David P. Corey[3]

[1] Department of Chemistry and Biochemistry, The Ohio State University, Columbus OH 43210, USA.

[2] Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA.

[3] Howard Hughes Medical Institute and Department of Neurobiology, Harvard Medical School, Boston, MA 02115, USA.

Keywords: adhesion, neuronal connectivity, hearing, mechanotransduction, cancer, brain development

Abstract

Members of the cadherin superfamily of proteins are involved in diverse biological processes such as morphogenesis, sound transduction, and neuronal connectivity. Key to cadherin function is their extracellular domain containing cadherin repeats, which can mediate interactions involved in adhesion and cell signaling. Recent cellular, biochemical, and structural studies have revealed that physical interaction among cadherins is more complex than originally thought. Here we review work on new cadherin complexes and discuss how the classification of the mammalian family can be used to search for additional cadherin complexes. We also highlight some of the challenges in cadherin research, namely, the characterization of a cadherin connectome in biochemical and structural terms, as well as the elucidation of molecular mechanisms underlying the functional diversity of non-classical cadherins *in vivo*.

Cadherin Superfamily

The classical cadherins were discovered as glycoproteins that mediate calcium-dependent cell-cell adhesion in early vertebrate embryo development and epithelial tissues [1-6]. Subsequent identification and characterization of multiple members of the superfamily revealed function beyond adhesion, depicting them as cell-surface receptors involved in signaling [7-10], mechanotransduction [11,12], and brain morphogenesis and wiring [13-15]. Cadherins, referred to here as all members of the superfamily (including protocadherins [16-17] and others), have also been found in multiple species and categorized in several subfamilies across phyla [18-21].

Characterization of the superfamily has been difficult due to the family's many members and their functional diversity. The classical cadherins involved in cell-cell adhesion have been studied in depth and molecular mechanisms underlying their biological function have been probed extensively in tetrapods [22]. However, less is known about the function of non-classical members and about members in other species [13,15,23-27].

Here we start with a brief, structurally-based introduction to the mammalian superfamily and its classification. We then focus on new molecular mechanisms involved in interactions among cadherin superfamily members in tetrapods and fish. We discuss the emergent view that cadherins might function as heterotypic receptor complexes (formed by more than one member of the superfamily), and on how classification of the family, sequence similarities, and structural relationships may help to identify new interactions among family members.

Cadherin Architecture and Structural Diversity

The hallmark of cadherin proteins is the cadherin extracellular repeat (EC) [28-30] (Figure 1), with cadherins defined as having at least two tandem EC repeats in their extracellular domain. In addition, cadherins usually have a transmembrane and a C-terminal cytoplasmic domain (Figure 1c). Each EC repeat consists of approximately 100 amino acids sharing a common fold, but not identical sequence, with seven β strands in a "Greek-key" motif (Figure 1a). The repeats arrange in series and linker regions feature multiple highly conserved amino acids that coordinate calcium ions in three calcium-binding sites (Figure 1b). The sequence motifs involved in calcium binding have enabled the identification of over 100 superfamily members in humans, and more throughout multiple species [31,27].

Classical cadherin proteins involved in calcium-dependent cell-cell adhesion in mammals, such as CDH1 and CDH2 (E- and N-cadherin; human nomenclature used unless otherwise noted), have been the most studied and serve as the archetypical members of the superfamily [32]. These cadherins have five extracellular EC repeats labeled EC1 to EC5, starting from the repeat most distal to the membrane at the N-terminus of the protein (Figure 1c). Cleavage of a N-terminal "prodomain" is required to enable their adhesive function [33], which is achieved by *trans* interactions, generally homotypic, between extracellular cadherin domains protruding from adjacent cells [34,35] (Figure 1e). Classical cadherins may also interact in parallel with partner molecules from the same cell through *cis* interactions [36-38], with *trans* bond formation preceding *cis* complex formation [22] (Figure 1d). The cytoplasmic domain of classical cadherins binds to catenins (p120 and β), which in turn regulate stability of the extracellular cadherin bond, endocytosis, and interaction between cadherins and the cytoskeleton through α catenin and other proteins [39-44].

Similarly, other cadherins are involved in *trans* and *cis* interactions, and use their cytoplasmic domain to bind to regulatory proteins (see Box 1). However, non-classical cadherins display a more diverse set of cytoplasmic domains, and their extracellular regions have varying number of EC repeats (Table 1). For instance, CDH13 (T-cadherin) with five EC repeats, is unique as it lacks transmembrane and cytoplasmic domains and is anchored to the membrane through a glycophosphatidylinositol (GPI) moiety [45,46]. The non-classical cadherins CDH16 and CDH17 have seven EC repeats and very short cytoplasmic domains [47-49]. On the protocadherin side, three gene clusters (α, β, and γ) code for a large number of proteins (a total of ~60 in most mammalian species) each with six EC repeats and a single pass transmembrane domain. The α and γ clusters have a variable cytoplasmic subdomain concatenated to a constant cytoplasmic region, whereas PCDHβs lack the latter [16,17,24,50]. Other "non-clustered" protocadherins have signature sequences in their cytoplasmic domains [51-54] or feature very long extracellular domains containing up to 34 EC repeats (Table 1 [54]). Overall, members of the superfamily have distinct and varied structural features that are relevant for their sometimes poorly understood function. These features also segregate members in different subfamilies, hinting that classification of the family according to sequence and structure may provide insights into evolutionary and functional relationships. However, a consistent classification has not been straightforward.

Cytoplasmic-based Classification of Cadherins

Initial attempts to classify the cadherin family used functional criteria and cytoplasmic domain sequences, rather than full-length protein sequences. This approach permitted easy identification of

classical cadherins in different species having similar function and cytoplasmic domains, but varied extracellular domains. For instance, Drosophila E-cadherin (DE-cadherin [31,55]) is essential for epithelial adherens junction formation and binds Drosophila catenins, just as CDH1 functions in cell-cell adhesion and binds to mammalian catenins. However, DE-cadherin's extracellular domain has seven EC repeats (instead of five ECs in CDH1), as well as an epidermal-growth-factor-like domain, a laminin globular domain, and a primitive proteolytic site, all three domains absent in CDH1 and in all mammalian classical cadherins [19]. Thus classification according to function and properties of the cytoplasmic domain, which does place DE-cadherin and CDH1 together in the classical subfamily, emerged as a good strategy to sort the family.

In this scheme, cadherins are pragmatically segregated into three groups: classical cadherins (able to bind catenins [19,56]), desmosomal cadherins, and protocadherins (PCDHs), which have a variable number of EC repeats (other than five) and cytoplasmic domains distinct from those of the first two subfamilies (Figure 1c).

As new members of the family in tetrapods and other species were discovered, new subfamilies, some with a single member, were incorporated and classified [13] (Table 1). In this cyto-centric extended classification, the classical subfamilies in the human genome are divided into type I and type II, the desmosomal cadherins comprise the desmocolin and desmoglein subfamilies, and the clustered protocadherins are readily separated into α, β, and γ subfamilies (Box 2 and Figure II). Human non-clustered protocadherins are less clearly separated in subfamilies unless other "ad-hoc" criteria are taken into account, such as the length of their cytoplasmic and extracellular domains (7D subfamily) or the specific cytoplasmic sequence signatures that define subfamilies δ1 and δ2. Finally, unclassified members are loosely grouped into the ε subfamily. Although classification according to

features of cytoplasmic sequences is less effective for non-classical cadherins (Figure II), the relevance of most cadherin cytoplasmic domains in signaling and disease further validated this approach [13] (see Box 1). Such classification likely provides insights into common signaling pathways used by members within one species and perhaps across different species, but it is unclear whether it reflects accurate evolutionary relationships as well as *trans* and *cis* connectivity among cadherins.

EC1-based Classification of Cadherins.

In parallel to cyto-centric efforts to classify the cadherin family, an alternative approach was developed based on the functional relevance of repeat EC1 [18,57]. Early research on classical cadherins showed that *trans* interactions were calcium-dependent and strictly homotypic, i.e, members of the subfamily interact only with identical types. Using chimeric cadherins it was shown that the EC1 repeat (most distal from the membrane) is responsible for binding specificity [35]. Over the past two decades structural [58] and biophysical [59,60] studies have further demonstrated that *trans* interaction of mammalian classical cadherins is mediated by the N-terminal extracellular EC1 repeat [58,61,62]. These interactions involve the exchange of secondary structure elements (short segments of β-strand A in EC1) along with exchanged-docking of one (type I) or two (type II) tryptophan side chains at the very N-terminus of the mature protein (Figure 2). While there is some variation in the surfaces involved in homotypic interactions among type I and type II members, all vertebrate classical cadherins studied so far use the same basic mechanism [22]. Moreover, extracellular domains of classical cadherins can drive adhesion by themselves (without cytoplasmic domains [61]).

The functional relevance of the extracellular domain and the involvement of EC1 in signaling and related binding mechanisms suggested that the whole family could be classified by aligning the sequences of EC1 repeats alone [57]. This provided a way to avoid the difficulties in comparing family members with vastly different number of EC repeats or highly divergent cytoplasmic domain sequences. Remarkably, EC1-based alignments effectively segregate vertebrate cadherins in groups that are similar to those obtained using "ad-hoc" criteria along with cytoplasmic-based alignments [18,13,20] (Table 1). In this scheme, human cadherins naturally separate in groups that include type I, type II, desmosomal, and 7D cadherins within the newly defined C-1 family of the cadherin major branch. Similarly, the α, β, and γ groups of clustered protocadherins separate from δ1 and δ2 protocadherins in the Cr-1a branch of the cadherin related major branch.

However, the EC1-based classification differs in some aspects to that based on cytoplasmic properties. In the human superfamily, clustered protocadherins PCDHαc1, αc2, γc3, γc4, and γc5 feature distinct extracellular EC1 repeats that group them away from their corresponding α and γ branches [24,50,63] (defined by cytoplasmic-based alignments and gene structure). The EC1-based alignment also segregates PCDH10 away from the δ2 protocadherins [20]. In addition, some cytoplasmic-based alignments that use a subset of cadherin members [20,54] suggest that PCDH15 groups with CDHR1 (PCDH21), CDH23 with CDHR2 (PCDH24), and PCDH12 with CDHR5 (μPCDH). Cytoplasmic-based alignments using the whole set of human cadherins are less conclusive about these subgroups, while the EC1-based alignment suggests the existence of three well-defined groups: the first one formed by CDHR1, CDHR2, and CDH23 (Cr-2 subfamily) and two additional single-member groups including CDHR5 and PCDH15 each [20]. Members inconsistently classified are highlighted in red in Table 1.

Importantly, new subfamilies can be identified in the EC1-based classification when considering non-mammalian sequences of cadherins [20,21]. The C-2 family, including the CELSRs in mammals, also encompasses type III cadherins like Drosophila DN-cadherin, chicken Hz, and zebrafish and Xenopus cHz-like cadherins. Type IV cadherins like Drosophila shotgun (DE-cadherin) and cricket (Gryllus bimaculatus) Gb1-cadherin also belong to the C-2 family. Type III cadherins with 13 to 17 EC repeats and type IV cadherins with seven EC repeats are not present in mammals.

The EC1-based scheme is not perfect. Proteins such as CDH1 and DE-cadherin are grouped in different subfamilies although they perform similar tasks in different species. Therefore, function may not be segregated with subfamily in this classification, which hints at more general difficulties when establishing evolutionary and sequence-structure-function relationships within a family of proteins across phyla [64]. Clearly, even close homologues may have different cellular functions and biochemical properties. Conversely, proteins that differ widely in sequence and structure may carry out similar functions in different species. Thus, evolutionary studies and sequence analyses of cadherins must be performed in the context of functional and biochemical data for members of protein families across multiple species [64].

Another shortfall of the EC1-based classification is that some cadherins use EC repeats other than EC1 to perform their function in adhesion and signaling. This could be particularly relevant for long cadherins with unusual calcium binding sites that may adopt globular shapes [65] thereby using binding mechanisms that resemble those used by the Drosophila DSCAMs [66].

Towards Cadherin Connectomics.

The EC1-based classification that emerged from work on classical cadherins remains attractive despite the pitfalls described above. It provides a simple way to deal with diverse multi-domain cadherins, while at the same time providing insights into EC1-based interaction mechanisms for cadherins. New binding mechanisms in classical and non-classical members of the family, however, suggest that the EC1-based homotypic interaction paradigm may not be applicable for all cadherins, and that the EC1-based classification should be updated. We therefore continue reviewing these new binding mechanisms observed for classical cadherins, atypical cadherins CDH23 and PCDH15, and clustered protocadherins, and suggest a new grouping based on the sequences of the first three EC repeats (Figure IIb), as well as new ways of finding potential interacting partners (Figure IIc,d and e; Box 2).

New Homotypic Trans Interaction Mechanisms Involve EC1 and EC2 Repeats

The involvement of multiple EC repeats in cell-cell adhesion mediated by classical cadherins has been highly debated over the past decade [67]. Biophysical and cell-based assays have suggested that classical cadherins could interdigitate their extracellular domains. In this arrangement, repeats EC2 to EC5 would also form part of the adhesive *trans* interface. However, single molecule FRET experiments [59,60] and cryo-electron tomography of desmosomes [68,69] indicate that interactions are tip-to-tip (EC1 to EC1), and reminiscent of the arrangements seen in crystals of entire extracellular domains of classical cadherins [61]. These studies further validate the approach of using EC1 sequences to classify the superfamily.

Interestingly, CDH13 (T-cadherin), a non-classical member of the cadherin family that lacks the EC1 features required for the *trans* EC1-to-EC1 strand-swapping interaction ([22], Figure 2a), does mediate *trans* homotypic cell-cell adhesion [45,46]. This unusual cadherin does not have a transmembrane or cytoplasmic domain, but it is linked to the plasma membrane through a GPI anchor. Moreover, crystallographic structures of CDH13 EC1 and EC2 repeats revealed a new mode of *trans* homotypic interaction that involves both its EC1 *and* EC2 repeats in a so-called "X-dimer" conformation [70] (Figure 2b). A series of *in vitro* biochemical and cell-based assays confirmed the X-dimer interface for CDH13, which was found to mediate robust aggregation of CHO-cells and to mediate inhibition of neurite outgrowth [70].

The X-dimer arrangement had been identified in previous X-ray structures of a mutant CDH1 [71], but had been regarded as either a candidate for mediating *cis* interactions among cadherins [71,38,72] or a crystal packing artifact (see discussion in [73]). Additional structural analyses of mutated classical cadherins and CDH13 [70,73] suggest that the X-dimer is a transient state used by *all* classical cadherins. This state, which is consistent with a second bonded conformation observed in single-molecule FRET experiments [60] and involves repeat EC2, may facilitate the subsequent exchange of β-strands seen in the EC1-to-EC1 strand swapped dimers [73] (Figure 2c). Thus, EC2 becomes relevant for homotypic *trans* interactions, and should be considered in EC-based classifications of the superfamily.

A New Trans Heterotypic Complex also Involves EC1 and EC2 Repeats

Just as the involvement of multiple EC repeats in cadherin *trans* interactions has been disputed, the existence of physiologically relevant heterotypic *trans* complexes of classical cadherins has been equally controversial. While the debate continues for classical cadherins, recent work has unequivocally shown that two non-classical cadherins, mouse Cdh23 and Pcdh15, do form a functional heterotypic complex that also involves both repeats EC1 *and* EC2.

The non-classical Cdh23 and Pcdh15 mouse proteins form hair-cell tip links [74-76], fine filaments essential for inner-ear mechanotransduction [12]. Tip links assemble and regenerate in a calcium-dependent manner, are constantly under tension, and connect adjacent stereocilia of the same cell (Figure 3a). The Cdh23 and Pcdh15 proteins, with 27 and 11 EC repeats, respectively, are products of deafness genes [76], are localized to tip links by antibody labeling, and their predicted length matches the length of the tip link, suggesting a tip-to-tip interaction [75,77] (Figure 3a,b). Moreover, binding experiments show robust heterotypic interaction *in vitro* [75,78].

The structures of the heterotypic complex formed by the interacting tips of mouse Cdh23 and Pcdh15 reveal that these proteins do not use any of the binding mechanisms observed for classical cadherins [78-80]. Instead, the EC1+2 protomers of each protein engage in an antiparallel extended "handshake", with both repeats contributing to the interface [79] (Figure 3c). Consistent with this structure, an arginine residue mutated to glycine in individuals with inherited deafness [81] is located in the handshake interface (Figure 3d). The structure predicts that the arginine-to-glycine mutation would disrupt molecular interactions thereby impairing binding, and this was confirmed by size exclusion chromatography and calorimetry experiments [79], as well as by prior binding assays [75] and *ex vivo* functional tests [77].

The handshake interface is possible due to special structural features of both Cdh23 and Pcdh15 protomers. In both cases the N-terminal strand of EC1 is unlike that of classical cadherins: it extends towards the top of the protomer where it is tucked and secured by a novel calcium binding site in Cdh23 [78,79] and by a disulfide bond in Pcdh15 [79] (Figure 3c,e,f). Both EC1 protomers present bulges that fit in the narrower inter-repeat linker region of the adjacent molecules, thereby facilitating the overlap. Interestingly, some mouse and human PCDH15 isoforms have modified N-terminal sequences, suggesting that PCDH15 across species may form new types of complexes with itself or with other members of the cadherin family. Moreover, recent experiments suggest that immature tip links might be transiently made of mouse Pcdh15 tetramers, without any Cdh23 molecules [82]. The multiple isoforms of mouse Pcdh15 [74] may provide a pseudo-heterotypic antiparallel bond responsible for the all-Pcdh15 immature tip links observed in [82].

The consequences of preventing the Cdh23 – Pcdh15 handshake interaction *in vivo* were determined using the *Noddy* mouse model [83]. *Noddy* mice carry an isoleucine to asparagine mutation at position 108, also located in the handshake interface. These mice completely lack inner-ear function. *In vitro* experiments show that this mutation impairs binding of Pcdh15 EC1+2 I108N to Cdh23 EC1+2, without preventing proper folding of the protomer. *In vivo* localization is also unaffected, indicating that impaired binding between the two proteins solely causes that phenotype. Thus the heterotypic handshake interaction, extensively validated *in vitro, ex vivo,* and *in vivo*, is essential for inner-ear mechanotransduction [79,83].

The structure of the heterotypic cadherin bond formed by Cdh23 and Pcdh15 indicates that EC2 repeats can also be involved in *trans* binding, as in the homotypic classical X-dimer interface. Thus, the EC1-based classification should be extended to include other repeats. The new classification could

be used to search for similar heterotypic binding pairs, as it is now clear that other cadherins (like the clustered protocadherins) form heterotypic complexes [15].

Protocadherin Complexes Involve Multiple Family Members and EC Repeats

The clustered protocadherins form the largest cadherin subfamily with 53 members in humans [16,17,24,50]. Many members have been studied individually, or collectively by antibody targeting of the constant domains in the cytoplasmic domains of α or γ protocadherins. Clustered protocadherins tested so far do not share the robust adhesive properties of classical cadherins, but rather cluster together in *cis* to form signaling receptors that mediate neuronal recognition [14,17,84] and survival [85-88] through presumably weak *trans* interactions.

Several issues have delayed the characterization of interactions among clustered protocadherins. First, bead-based binding experiments fail to show interaction among family members [15]. In addition, cell-based assays did not show strong calcium-dependent cell aggregation activity for multiple members of the subfamily [16,89-94]. Lastly, interpretation of these results might have been confounded by endogenous expression of γ protocadherins [95].

To overcome these difficulties and to probe interactions among mouse Pcdhγ complexes, non-adherent K562 cells, which lack endogenous expression of classical cadherins and γ protocadherins have been used [95-97]. These experiments revealed homotypic *trans* interaction for Pcdhγa3, b2, and c3 [95]. The interaction mediated by Pcdhγa3 was found to be only partially calcium-dependent and

required repeat EC1. Additional experiments with Pcdhγa10, a12, b1, b6 confirmed strict homotypic *trans* interactions. Remarkably, chimeric constructs show that the specificity of these interactions is governed by repeats EC2 and EC3 [95], not EC1 (Figure 3g). This is consistent with positively selected codon positions found mostly in EC2 and EC3 repeats when analyzing EC1-3 sequences [24].

Immunoprecipitation experiments also suggest that members of the Pcdhγ subfamily form tetrameric complexes in *cis* [95] (Figure 3g). The number of *cis* heterotypic tetramers that can be formed and made available for *trans* homotypic interaction is large enough to provide a rich repertoire of unique receptors that can form the basis for a neuronal recognition code [84,95]. Interestingly, recent experiments suggests that members of the α, β, and γ subfamilies interact (directly or indirectly) with each other, although not necessarily to mediate adhesion [98-102]. Thus, the cadherin code for neuronal recognition could be greatly extended.

Proteomic analysis using mouse proteins also suggest that members of the α and γ subfamilies interact with the classical cadherin Cdh2 [98], but likely in an indirect way [99]. Similarly, members of the γ subfamily may interact with other members of the superfamily, like Cdh4 and Pcdh17 [98], while members of the δ1, δ2, and ε subfamilies of non-clustered protocadherins may also form heterotypic complexes [103-106] (Figure 3g,h).

For instance, Cdh2 and Pcdh19 functionally cooperate with each other in zebrafish brain development [105], and their physical interaction has been the best characterized so far [106]. While zebrafish Pcdh19 does not mediate adhesion on its own, and does not interact in *trans* with zebrafish Cdh2, it is responsible for the *trans* adhesive interaction of the heterotypic Cdh2-Pcdh19 complex. Cdh2 seems to enable the adhesion driven by Pcdh19, suggesting a paradigm-shifting view in which

classical cadherins, at least in zebrafish, act as regulators of δ2 protocadherin-mediated adhesion [15,106] (Figure 3h).

Overall, these results show an apparent heterotypic promiscuity among non-classical cadherins interacting in *cis* (particularly for the clustered protocadherins), while *trans* interactions that involve repeats EC1, EC2, *and* EC3 are strictly homotypic, at least for the γ subfamily. It is apparent that involvement of some cadherins in signaling and cell recognition requires a transient and perhaps weak interaction, while in other cases, a strong bond may serve for concomitant signaling and adhesive functions. Regardless of their role in signaling or in more stable adhesive contacts, most of the cadherin interactions described so far involve repeats that go beyond repeat EC1, and most have not been complemented with structures. Moreover, there is a lack of a systematic exploration of the possible complexes that can be formed by superfamily members.

Concluding Remarks

The research summarized above has revealed functional and structural diversity among cadherins. Overall, these results highlight the need to abandon the narrow view of homotypic adhesive interactions mediated by EC1 for all cadherins and explore function beyond the classical paradigm (see Outstanding Questions). Many cadherins use alternate binding mechanisms that involve both homotypic and heterotypic interactions mediated by multiple EC repeats. These interactions may play a role in adhesion, signaling, or both. Together these results also suggest that the EC1-based classification should be extended to incorporate at least EC2 and EC3 repeats. Evolutionary relationships are unlikely to be captured by such classification, as they require more comprehensive

analyses of the full-length protein sequences and of the biochemical properties of both intra and extracellular domains. However, we propose that EC-based classifications may serve a different purpose: to identify families and subfamilies, within a species, that share distinct binding mechanisms, thereby providing roadmaps to probe and build species-specific cadherin connectomes. For instance, rearrangement of the human cadherin superfamily using sequences covering EC1, EC2, and EC3 repeats reveals interesting relationships and general criteria that can be used to predict interaction candidates (Box 2). As a test, we propose two strong complex candidates for heterotypic interactions (FAT4/FAT3 and CDHR2/CDHR5), among many other possibilities that need to be systematically explored, as recently done for immunoglobulin and LRR proteins in Drosophila [107].

Ultimately, understanding how cadherins interact with each other will serve to design molecular handles that can control their function *in vivo*, perhaps to modify development and morphogenesis [108], or alter neuronal connectivity in a controlled fashion [109]. This may provide essential tools to probe the functional connectome of the brain [110].

Outstanding Questions BOX

The molecular mechanisms underlying function of cadherins have been elucidated in great detail for some members of the superfamily. However, multiple challenges need to be addressed to answer outstanding questions and to probe function of many more cadherins for which data are scarce. For instance, *ex vivo* and *in vitro* assays to test cadherin interactions often give confounding results and a decade of trials suggests that five conditions must be satisfied in *ex vivo* cell-based binding assays: 1) experiments should show that endogenous expression of all cadherins is limited and is not

interfering with the interactions being probed; 2) induced expression of cadherins and other proteins upon transfection must be checked as well; 3) control of surface expression by endocytosis and cleavage through metalloproteases must be considered when interpreting *ex vivo* results; 4) interactions with other non-cadherin proteins [91,111], not covered in this review, should be considered and probed; 5) *in vivo* and *ex vivo* interactions must be recapitulated *in vitro* to identify all components of the complexes and differentiate direct vs. indirect interactions. Similarly, *in vitro* bead-based binding assays should take into account the lack of cytoplasmic domains and *cis* partners that might be required to establish an interaction. In addition, the role of mechanical forces as modulators of binding affinity, as well as isoform diversity and glycosylation must be taken into account in the context of testing functional cadherin interactions.

Additional challenges arise because *ex vivo* and *in vitro* binding assays are often carried out with shorter recombinant versions of the wild-type proteins. This may hinder the elucidation of interaction mechanisms that involve other repeats, especially for long cadherins, which may lack an obvious EC1 repeat or which may fold in globular shapes [65]. Site-directed mutagenesis is also often used to probe molecular interaction mechanisms, but these mutations may have pleotropic effects, for example affecting folding or calcium binding affinity. Impaired calcium binding may have subtle effects that prevent binding, even under saturating calcium concentrations.

Addressing these issues and challenges systematically may provide a way to answer several open questions: 1) What is the exact stoichiometry and composition of cadherin complexes, especially those formed by clustered protocadherins? 2) What are the molecular mechanisms underlying *cis* interactions, and are there new ways to form *trans* complexes? 3) What determines homotypic and heterotypic specificities (or lack thereof)? 4) Are all signaling interactions among cadherins intrinsically weak? 5) How do the extracellular domains of cadherins control the binding of signaling

molecules to their own cytoplasmic tails? 6) How do long cadherin extracellular domains arrange on the membrane surface?

BOX 1 Family members and signaling

Since the discovery of interactions between classical cadherins and the catenins involved in gene transcription, it has been clear that cadherins may function in both cellular adhesion and intracellular signaling. For instance, while the classical cadherin extracellular domain provides an adhesive bond and acts as a calcium sensor, differential cadherin affinities in homotypic binding mediated by EC1 modulate GTPase signaling [112]. On the clustered protocadherin side, the α and γ subfamilies have been implicated in neuronal survival by signaling through binding of their cytoplasmic domains to kinases [17,101] (Figure I).

Other "non-clustered" protocadherins have cytoplasmic signature sequences that determine cytoplasmic signaling partners (CM1, CM2, and CM3 sequences for δ1 protocadherins and CM1 and CM2 for δ2 protocadherins [51-54]). δ1 protocadherins interact with protein phosphatase-1α (PP1α), the histone-regulating TATA-binding protein-associated factor-1 TAF1 (Pcdh7), or β-catenin (PCDH11Y). Similarly, the δ2 protocadherins interact with the serine-threonine kinase TAO2β (pcdh8), the actin regulatory complex Nap1/WAVE1 (pcdh10), and mouse disabled-1 mDab1 (pcdh18; [54] and references therein; Figure I). In addition, interaction with the WAVE regulatory complex (WRC) is mediated by WRC interacting sequences (WIRS) present in the α and δ2

subfamilies [113]. Overall, these cytoplasmic partners highlight the role of protocadherins in intracellular signaling.

Other family members, often featuring long extracellular domains and loosely grouped into the protocadherin ε subfamily (Table 1 [54]) have been implicated in signaling processes that involve interaction of their cytoplasmic domains with diverse intracellular partners [114]. Among them, dachsous 1 (Dchs1) and Fat4 play a role in brain morphogenesis and planar cell polarity in kidneys and inner ear ([114] and references therein). Fat4 interacts with MUPP1 and LIX/LIX1L. Fat1, involved in CNS and kidney development, binds to β-catenin, atrophins, Ena/VASP, HOMER, and Scribble ([113] and references therein). Fat3 has been implicated in neuronal morphology [115], but less is known about its cytoplasmic partners. Celsrs are important in planar cell polarity and neuronal morphology, but again little is known about their intracellular partners [116].

Two members of the ε subfamily stand out with rather unique functional features. PCDH15 and CDH23 are important for sensory perception [12] (hearing, balance, and sight, see main text). In hair cells of the inner ear, these cadherins participate directly in mechanotransduction by linking adjacent stereocilia from the same cell. In addition, their multiple intracellular partners (sans, whirlin, harmonin, and myosin VIIa) have been proposed to form a signaling protein network that is dysfunctional in Usher syndrome, a genetic disorder that causes deafness and blindness [117,118] (Figure I).

Despite the many cases in which cytoplasmic binding partners of cadherins have been identified, it still unclear how extracellular and intracellular domains work together to integrate and trigger corresponding signaling cascades.

BOX 2 Identifying New Cadherin Complexes

To predict interactions among cadherins the EC1-3 based classification is used (Table 1 and Figure IIb). Specific rules of engagement can be defined based on our current knowledge of cadherin interactions. For instance, homotypic and heterotypic *trans* interactions have been reported for members within subtrees defined by type I, type II, and desmosomal cadherins, but not across these subtrees. Similarly, FAT4 groups with DCHS1 and DCHS2, and *trans* heterotypic interactions between Fat4 and Dchs1 have been reported. Therefore, members of a given subtree in which there is at least one confirmed *trans* homotypic or heterotypic interaction, *may* similarly engage in heterotypic *trans* interactions with some of the other members of the same subtree. This "subtree identity" criterion is valid for the cases mentioned above, and makes specific predictions (Figure IIc).

As an example, given that zebrafish Pcdh19 mediates homotypic *trans* adhesive interactions (and assuming that this result is valid for mammals), the subtree criterion used with the EC1-3 based alignment predicts heterotypic *trans* adhesive interactions among PCDH10, PCDH17, PCDHαc2, PCDHγc4, and PCDHγc5 (which group with the PCDH19 subtree). Similarly, PCDH8, PCDH12, and PCDH18, as well as all δ1 protocadherins would form two groups with potential for *trans* homotypic and heterotypic contact formation (Figure IIc). Members of these groups do not have long N-termini, required for a handshake interaction, and lack the tryptophans required for a β–strand classical interaction, suggesting an X-dimer or a novel type of interaction (except for PCDH12 and the δ1 PCDH20, which do feature long N-termini and may use a handshake-like interaction instead).

The subtree identity criterion is not be applicable to clustered protocadherins, which do not seem to mediate adhesive *trans* interactions, and even weak *trans* interactions have been reported to be strictly homotypic (involving a single branch within the subtree). Similarly, CELSRs do not engage in heterotypic *trans* interactions.

Another plausible criterion arises when considering interactions among members of different subtrees. For instance CDH23, which groups with CDHR1 and CDHR2 (Cr-2), engages in *trans* heterotypic interactions with PCDH15, which groups along with CDHR3, CDHR4, CDHR5, CDH16, CDH17, FAT1, FAT2, and FAT3 (Cr-3; Figure IId). We define these subtrees as "siblings" and postulate that their members have the potential to engage in heterotypic contact. Consistently, CDHR2 and CDHR5, showing overlapping expression in the enterocyte brush border, are postulated to form a complex [119]. These two proteins share structural features important for handshake-like interactions (long N-termini, a putative CDH23-like calcium-binding site for CDHR2, and a putative PCDH15-like disulfide bond for CDHR5). Thus, this subset of cadherins may engage in handshake-like interactions. Similarly, interactions between CDH17 (Cr-3) and CDH1 (Type I) [120,121] define another set of potential interactions (Figures 3i and IId).

A refinement of these criteria may involve specific sequence motifs that facilitate known cadherin interactions. For instance, the PCDH15 "EV**R**IVVR" motif involved in the handshake with CDH23 is found in mouse Fat4 at the same location ("EV**R**VLVR") [79]. Similarly, the CDH23 "KV**N**I**Q**V" motif with interfacial residues is *identical* in EC1 of mouse Fat3. Thus, Fat4 and Fat3 may engage in a handshake-like interaction, and define another set of sibling subtrees.

Additional sibling subtrees can be defined when extending this criterion to *cis* heterotypic contacts. Pcdh19 from zebrafish and Pcdh8 from mouse form a *cis* complex with CDH2, defining two additional sets of sibling subtrees (δ2 and type I; δ2′ and type I) and thereby multiple potential complexes (Figure IIe). Similarly, the α, β and γ clustered protocadherins seem to engage promiscuously in *cis* heterotypic complexes. The underlying molecular mechanisms and specificity (if any) for these interactions remain unknown.

Acknowledgments

We thank members of the Sotomayor, Corey, and Gaudet laboratories, as well as A. Krol for valuable discussions. This work was supported by the NIH (K99/R00 DC012534-01 to MS and R01 DC02281 to DPC). DPC is an investigator of the Howard Hughes Medical Institute.

References

[1] Takeichi M (1977) Functional correlation between cell adhesive properties and some cell surface proteins. Journal of Cell Biology 75, 464-474.

[2] Hyafil F, Babinet C, Jacob F (1981) Cell-cell interactions in early embryogenesis: a molecular approach to the role of calcium. Cell 26, 447-454.

[3] Damsky C, Knudsen KA, Dorio RJ, Buck CA (1981) Manipulation of cell-cell and cell-substratum interactions in mouse mammary tumor epithelial cells using broad spectrum antisera. Journal of Cell Biology 89, 173-184.

[4] Bertolotti R, Rutishauser U, Edelman GM (1980) A cell surface molecule involved in aggregation of embryonic liver cells. Proceedings of the National Academy of Sciences, USA 77, 4831-4835.

[5] Takeichi M (1990) Cadherins: A molecular family important in selective cell-cell adhesion. Annu. Rev. Biochem. 59, 237-252.

[6] Kemler R (1992) Classical cadherins. Semin. Cell Biol. 3, 149-155.

[7] Suzuki ST (1996) Structural and functional diversity of cadherin superfamily: are new members of cadherin superfamily involved in signal transduction pathway? J. Cell Biochem. 61, 531-542.

[8] Vleminckx K, Kemler R (1999) Cadherins and tissue formation: integrating adhesion and signaling. Bioessays 21, 211-220.

[9] Wheelock MJ, Johnson KR (2003) Cadherin-mediated cellular signaling. Curr. Opin. Cell Biol. 15, 509-514.

[10] Saito M, Tucker DK, Kohlhorst D, Niessen C, Kowalczyk AP (2012) Classical and desmosomal cadherins at a glance. J. Cell Sci. 125, 2547-2552.

[11] Leckband DE, le Duc Q, Wang N, de Rooij J (2011) Mechanotransduction at cadherin mediated adhesions. Curr. Opin. Cell Biol. 23, 523-530.

[12] Gillespie PG, Muller U (2009) Mechanotransduction by hair cells: Models, molecules, and mechanisms. Cell 139, 33-44.

[13] Hirano S, Takeichi M (2012) Cadherins in brain morphogenesis and wiring. Physiol. Rev. 92, 597-634.

[14] Lefebvre JL, Kostadinov D, Chen WV, Maniatis T, Sanes JR (2012) Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. Nature 488, 517-521.

[15] Weiner JA, Jontes JD (2013) Protocadherins, not prototypical: a complex tale of their interactions, expression, and functions. Front. Mol. Neurosci. 6, 1-4.

[16] Sano K, et al. (1993) Protocadherins: a large family of cadherin-related molecules in central nervous system. EMBO J. 12, 2249-2256.

[17] Kohmura N, et al. (1998) Diversity revealed by a novel family of cadherins expressed in neurons at a synaptic complex. Neuron 20, 1137-1151.

[18] Nollet F, Kools P, van Roy F (2000) Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members. Journal of Molecular Biology 299, 551-572.

[19] Oda H, Takeichi M (2011) Evolution: structural and functional diversity of cadherin at the adherens junction. J. Cell Biol. 193, 1137-1146.

[20] Hulpiau P, van Roy F (2009) Molecular evolution of the cadherin superfamily. The International Journal of Biochemistry and Cell Biology 41, 349-369.

[21] Hulpiau P, Gul IS, van Roy F (2013) New insights into the evolution of metazoan cadherins and catenins. Prog. Mol. Biol. Transl. Sci. 116, 71-94.

[22] Brasch J, Harrison OJ, Honig B, Shapiro L (2012) Thinking outside the cell: how cadherins drive adhesion. Trends Cell Biol. 22, 299-310.

[23] Tada MN, et al. (2004) Genomic organization and transcripts of the zebrafish protocadherin genes. Gene 340, 197-211.

[24] Wu Q (2005) Comparative genomics and diversifying selection of the clustered vertebrate protocadherin genes. Genetics 169, 2179-2188.

[25] Yu WP, et al. (2008) Elephant shark sequence reveals unique insights into the evolutionary history of vertebrate genes: A comparative analysis of the protocadherin cluster. Proceedings of the National Academy of Sciences, USA 105, 3819-3824.

[26] Jiang XJ, Li S, Ravi V, Venkatesh B, Yu WP (2009) Identification and comparative analysis of the protocadherin cluster in a reptile, the green anole lizard. PLoS One 4, e7614.

[27] Nichols SA, Roberts BW, Richter DJ, Fairclough SR, King N (2012) Origin of metazoan cadherin diversity

and the antiquity of the classical cadherin/β-catenin complex. Proceedings of the National Academy of Sciences, USA 109, 13046-13051.

[28] Ringwald M, et al. (1987) The structure of cell adhesion molecule uvomorulin. insights into the molecular mechanism of Ca2+-dependent cell adhesion. EMBO J 6, 3647-3653.

[29] Overduin M, et al. (1995) Solution structure of the epithelial cadherin domain responsible for selective cell adhesion. Science 267, 386-389.

[30] Shapiro L, et al. (1995) Structural basis of cell-cell adhesion by cadherins. Nature 374, 327-336.

[31] Oda H, Uemura T, Harada Y, Iwai Y, Takeichi M (1994) A Drosophila homolog of cadherin associated with armadillo and essential for embryonic cell-cell adhesion. Dev. Biol. 165, 716-726.

[32] van Roy F, Berx G (2008) The cell-cell adhesion molecule E-cadherin. Cell. Mol. Life Sci. 65, 3756-3788.

[33] Ozawa M, Kemler R (1990) Correct proteolytic cleavage is required for the cell adhesive function of uvomorulin. Journal of Cell Biology 111, 1645-1650.

[34] Nagafuchi A, Shirayoshi Y, Okazaki K, Yasuda K, Takeichi M (1987) Transformation of cell adhesion properties by exogenously introduced E-cadherin cDNA. Nature 329, 341-343.

[35] Nose A, Tsuji K, Takeichi M (1990) Localization of specificity determining sites in cadherin cell adhesion molecules. Cell 61, 147-155.

[36] Brieher WM, Yap AS, Gumbiner BM (1996) Lateral dimerization is required for the homophilic binding activity of C-cadherin. J. Cell Biol 135, 487-496.

[37] Takeda H, Shimoyama Y, Nagafuchi A, Hirohashi S (1999) E-cadherin functions as a cis-dimer at the cell-cell adhesive interface in vivo. Nature Structural Biology 6, 310-312.

[38] Pertz O, et al. (1999) A new crystal structure, Ca2+ dependence and mutational analysis reveal molecular details of E-cadherin homoassociation. EMBO Journal 18, 1738-1747.

[39] Nagafuchi A, Takeichi M (1988) Cell binding function of E-cadherin is regulated by the cytoplasmic domain. EMBO J 7, 3679-3684.

[40] Ozawa M, Baribault H, Kemler R (1989) The cytoplasmic domain of the cell adhesion molecule uvomorulin associates with three independent proteins structurally related in different species. EMBO Journal 8, 1711-1717.

[41] Abe K, Takeichi M (2008) EPLIN mediates linkage of the cadherin catenin complex to F-actin and stabilizes the circumferential actin belt. Proceedings of the National Academy of Sciences, USA 105, 13-19.

[42] Pokutta S, Drees F, Yamada S, Nelson WJ, Weis WI (2008) Biochemical and structural analysis of alpha-catenin in cell-cell contacts. Biochem. Soc. Trans. 36, 141-147.

[43] Shapiro L, Weis WI (2009) Structure and biochemistry of cadherins and catenins. Cold Spring Harb. Perspect. Biol. 1, a003053.

[44] Hong S, Troyanovsky RB TS (2013) Binding to F-actin guides cadherin cluster assembly, stability, and movement. J. Cell Biol. 201, 131-143.

[45] Ranscht B, Dours-Zimmermann MT (1991) T-cadherin, a novel cadherin cell adhesion molecule in the

nervous system lacks the conserved cytoplasmic region. Neuron 7, 391-402.

[46] Vestal DJ, Ranscht B (1992) Glycosyl phosphatidylinositol-anchored T-cadherin mediates calcium-dependent, homophilic cell adhesion. J. Cell Biol. 119, 451-461.

[47] Thomson RB, et al. (1995) Isolation and cDNA cloning of Ksp-cadherin, a novel kidney specific member of the cadherin multigene family. J. Biol. Chem. 270, 17594-17601.

[48] Berndorff D, et al. (1994) Liver-intestine cadherin: molecular cloning and characterization of a novel Ca(2+)-dependent cell adhesion molecule expressed in liver and intestine. J. Cell Biol. 125, 1353-1369.

[49] Wendeler MW, Jung R, Himmelbauer H, Gessner R (2006) Unique gene structure and paralogy define the 7D-cadherin family. Cell Mol. Life Sci. 63, 1564-1573.

[50] Wu Q, Maniatis T (1999) A striking organization of a large family of human neural cadherin like cell adhesion genes. Cell 97, 779-790.

[51] Wolverton T, Lalande M (2001) Identification and characterization of three members of a novel subclass of protocadherins. Genomics 76, 66-72.

[52] Vanhalst K, Kools P, Staes K, van Roy F, Redies C (2005) delta-protocadherins: a gene family expressed differentially in the mouse brain. Cell Mol. Life Sci. 62, 1247-1259.

[53] Redies C, Vanhalst K, Roy F (2005) delta-protocadherins: unique structures and functions. Cell Mol. Life Sci. 62, 2840-2852.

[54] Kim S, Yasuda S, Tanaka H, Yamagata K, Kim H (2011) Non-clustered protocadherin. Cell Adh. Migr. 5,

97-105.

[55] Woods DF, Wu JW, Bryant PJ (1997) Localization of proteins to the apico-lateral junctions of Drosophila epithelia. Dev. Genet. 20, 111-118.

[56] Takeichi M (1995) Morphogenetic roles of classic cadherins. Current Opinion in Cell Biology 7, 619-627.

[57] Pouliot Y (1992) Phylogenetic analysis of the cadherin superfamily. Bioessays 14, 743-748.

[58] Boggon TJ, et al. (2002) C-cadherin ectodomain structure and implications for cell adhesion mechanisms. Science 296, 1308-1313.

[59] Zhang Y, Sivasankar S, Nelson WJ, Chu S (2009) Resolving cadherin interactions and binding cooperativity at the single-molecule level. Proceedings of the National Academy of Sciences, USA 106, 109-114.

[60] Sivasankar S, Zhang Y, Nelson WJ, Chu S (2009) Characterizing the initial encounter complex in cadherin adhesion. Structure 17, 1075-1081.

[61] Harrison OJ, et al. (2011) The extracellular architecture of adherens junctions revealed by crystal structures of type I cadherins. Structure 19, 244-256.

[62] Patel SD, et al. (2006) Type II cadherin ectodomain structures: Implications for classical cadherin specificity. Cell 124, 1255-1268.

[63] Takei Y, et al. (2001) Two novel CNRs from the CNR gene cluster have molecular features distinct from those of CNR1 to 8. Genomics 72, 321-330.

[64] Dickinson DJ, Weis WI, Nelson WJ (2011) Protein evolution in cell and tissue development: going beyond sequence and transcriptional analysis. Dev. Cell 21, 32-34.

[65] Jin X, et al. (2012) Crystal structures of Drosophila N-cadherin ectodomain regions reveal a widely used class of Ca2+-free interdomain linkers. Proceedings of the National Academy of Sciences, USA 109, E127-134.

[66] Sawaya MR, et al. (2008) A double S shape provides the structural basis for the extraordinary binding speciffcity of Dscam isoforms. Cell 134, 1007-1018.

[67] Leckband D, Prakasam A (2006) Mechanism and dynamics of cadherin adhesion. Annual Review of Biomedical Engineering 8, 259-287.

[68] He W, Cowin P, Stokes DL (2003) Untangling desmosomal knots with electron tomography. Science 302, 109-113.

[69] Al-Amoudi A, Diez DC, Betts MJ, Frangakis AS (2007) The molecular architecture of cadherins in native epidermal desmosomes. Nature 450, 832-837.

[70] Ciatto C, et al. (2010) T-cadherin structures reveal a novel adhesive binding mechanism. Nat. Struct. Mol. Biol. 17, 339-347.

[71] Nagar B, Overdulin M, Ikura M, Rini J (1996) Structural basis of calcium-induced E-cadherin rigidification and dimerization. Nature 380, 360-364.

[72] Haussinger D, et al. (2004) Proteolytic E-cadherin activation followed by solution NMR and X-ray crystallography. EMBO Journal 23, 1699-1708.

[73] Harrison OJ, et al. (2010) Two-step adhesive binding by classical cadherins. NSMB 17, 348-357.

[74] Ahmed ZM, et al. (2006) The tip-link antigen, a protein associated with the transduction complex of sensory hair cells, is protocadherin-15. Journal of Neuroscience 26, 7022-7034.

[75] Kazmierczak P, et al. (2007) Cadherin 23 and protocadherin 15 interact to form tip-link filaments in sensory hair cells. Nature 449, 87-91.

[76] Corey, D. P. (2007) Stringing the fiddle: the inner ear's two-part invention. Nature Neuroscience 10:1232-1233.

[77] Lelli A, Kazmierczak P, Kawashima Y, Muller U, Holt JR (2010) Development and regeneration of sensory transduction in auditory hair cells requires functional interaction between cadherin-23 and protocadherin-15. J. Neurosci. 30, 11259-11269.

[78] Elledge HM, et al. (2010) Structure of te N terminus of cadherin 23 reveals a new adhesion mechanism for a subset of cadherin superfamily members. Proceedings of the National Academy of Sciences, USA 107, 10708-100712.

[79] Sotomayor M, Weihofen WA, Gaudet R, Corey DP (2012) Structure of a force-conveying cadherin bond essential for inner-ear mechanotransduction. Nature 492, 128-132.

[80] Sotomayor M, Weihofen WA, Gaudet RG, Corey DP (2010) Structural determinants of cadherin-23 function in hearing and deafness. Neuron 66, 85-100.

[81] Ahmed ZM, et al. (2008) Gene structure and mutant alleles of PCDH15: nonsyndromic deafness DFNB23 and type 1 Usher syndrome. Hum. Genet. 124, 215-223.

[82] Indzhykulian AA, et al. (2013) Molecular remodeling of tip links underlies mechanosensory regeneration in auditory hair cells. PLoS Biology e1001583.

[83] Geng R, et al. (2013) Noddy, a mouse harboring a missense mutation in protocadherin-15, reveals the impact of disrupting a critical interaction site between tip-link cadherins in inner ear hair cells. J. Neurosci. 33, 4395-4404.

[84] Yagi T (2012) Molecular codes for neuronal individuality and cell assembly in the brain. Front. Mol. Neurosci 5, 45.

[85] Wang X, et al. (2002) Gamma protocadherins are required for survival of spinal interneurons. Neuron 36, 843-854.

[86] Junghans D, Haas IG, Kemler R (2005) Mammalian cadherins and protocadherins: about cell death, synapses and processing. Curr. Opin. Cell Biol. 17, 446-452.

[87] Lefebvre JL, Zhang Y, Meister M, Wang X, Sanes JR (2008) gamma-Protocadherins regulate neuronal survival but are dispensable for circuit formation in retina. Development 135, 4141-4151.

[88] Chen WV, et al (2012) Functional significance of isoform diversification in the protocadherin gamma gene cluster. Neuron 75: 402-409.

[89] Sago H, et al. (1995) Cloning, expression, and chromosomal localization of a novel cadherin related protein, protocadherin-3. Genomics 29, 631-640.

[90] Triana-Baltzer GB, Blank M (2006) Cytoplasmic domain of protocadherin-alpha enhances homophilic

interactions and recognizes cytoskeletal elements. J. Neurobiol. 66, 393-407.

[91] Mutoh T, Hamada S, Senzaki K, Murata Y, Yagi T (2004) Cadherin-related neuronal receptor 1 (CNR1) has cell adhesion activity with beta1 integrin mediated through the RGD site of CNR1. Exp. Cell Res. 294, 494-508.

[92] Morishita H, et al. (2006) Structure of the cadherin-related neuronal receptor/protocadherin-α first extracellular cadherin domain reveals diversity across cadherin families. Journal of Biological Chemistry 281, 33650-33663.

[93] Frank M, et al. (2005) Diferential expression of individual gamma-protocadherins during mouse brain development. Mol Cell Neurosci 29, 603-616.

[94] Fernandez-Monreal M, Kang S, Phillips GR (2009) Gamma-protocadherin homophilic interaction and intracellular trafficking is controlled by the cytoplasmic domain in neurons. Mol. Cell Neurosci. 40, 344-353.

[95] Schreiner D, Weiner JA (2010) Combinatorial homophilic interaction between gamma protocadherin multimers greatly expands the molecular diversity of cell adhesion. Proceedings of the National Academy of Sciences, USA 107, 14893-14898.

[96] Reiss K, et al. (2006) Regulated ADAM10-dependent ectodomain shedding of gammaprotocadherin C3 modulates cell-cell adhesion. J. Biol. Chem. 281, 21735-21744.

[97] Ozawa M, Kemler R (1998) Altered cell adhesion activity by pervanadate due to the dissociation of alpha-catenin from the E-cadherin.catenin complex. J. Biol. Chem. 273, 6166-6170.

[98] Han MH, Lin C, Meng S, Wang X (2010) Proteomics analysis reveals overlapping functions of clustered protocadherins. Mol. Cell Proteomics 9, 71-83.

[99] Murata Y, Hamada S, Morishita H, Mutoh T, Yagi T (2004) Interaction with protocadheringamma regulates the cell surface expression of protocadherin-alpha. J. Biol. Chem. 279, 49508-49516.

[100] Bonn S, Seeburg PH, Schwarz MK (2007) Combinatorial expression of alpha- and gamma-protocadherins alters their presenilin-dependent processing. Mol. Cell Biol. 27, 4121-4132.

[101] Chen J, et al. (2009) alpha- and gamma-protocadherins negatively regulate PYK2. J. Biol. Chem. 284, 2880-2890.

[102] Biswas S, Emond MR, Jontes JD (2012) The clustered protocadherins Pcdhα and Pcdhγ form a heteromeric complex in zebrafish. Neuroscience 219, 280-289.

[103] Blevins CJ, Emond MR, Biswas S, Jontes JD (2011) Differential expression, alternative splicing, and adhesive properties of the zebrafish δ1-protocadherins. Neuroscience 199, 523-534.

[104] Yamagata K, et al. (1999) Arcadlin is a neural activity-regulated cadherin involved in long term potentiation. J. Biol. Chem. 274, 19473-19479.

[105] Biswas S, Emond MR, Jontes JD (2010) Protocadherin-19 and N-cadherin interact to control cell movements during anterior neurulation. J. Cell Biol. 191, 1029-1041.

[106] Emond MR, Biswas S, Blevins CJ, Jontes JD (2011) A complex of protocadherin-19 and N-cadherin mediates a novel mechanism of cell adhesion. J. Cell Biol. 195, 1115-1121.

[107] Ozka, E, et al. (2013) An Extracellular Interactome of Immunoglobulin and LRR Proteins Reveals

Receptor-Ligand Networks.

[108] Xiong F, et al. (2013) Specified neural progenitors sort to form sharp domains after noisy shh signaling. Cell 153, 550-561.

[109] Demireva EY, Shapiro LS, Jessell TM, Zampieri N (2011) Motor neuron position and topographic order imposed by β- and γ-catenin activities. Cell 147, 641-652.

[110] Morgan JL, Lichtman JW (2013) Why not connectomics? Nature Methods 10, 494-500.

[111] Li Y, et al. (2012) Molecular and functional interaction between protocadherin-γC5 and GABAA receptors. J. Neurosci. 32, 11780-11797.

[112] Tabdili H, et al. (2012) Cadherin point mutations alter cell sorting and modulate GTPase signaling. J. Cell Sci. 125, 3299-3309.

[113] Chen, B. et al. (2014) The WAVE regulatory complex links diverse receptors to the actin cytoskeleton. Cell 156:195-207.

[114] Sadeqzadeh E, de Bock CE, Thorne RF (2013) Sleeping giants: Emerging roles for the Fat cadherins in health and disease. Med. Res. Rev. 34, 190-221.

[115] Deans MR, et al. (2011) Control of neuronal morphology by the atypical cadherin Fat3. Neuron 71, 820-832.

[116] Boutin C, Goffinet AM, Tissir F (2012) Celsr1-3 cadherins in PCP and brain development. Curr. Top. Dev.

Biol. 101, 161-183.

[117] Friedman TB, Schultz JM, Ahmed Z, Tsilou ET, C. BC (2011) Usher syndrome: hearing loss with vision loss. Adv. Otorhinolaryngol. 70, 56-65.

[118] Bonnet C EAA (2012) Usher syndrome (sensorineural deafness and retinitis pigmentosa): pathogenesis, molecular diagnosis and therapeutic approaches. Curr. Opin. Neurol. 25, 42-49.

[119] McConnell RE, Benesh AE, Mao S, Tabb DL, Tyska MJ (2011) Proteomic analysis of the enterocyte brush border. Am. J. Physiol. Gastrointest. Liver. Physiol. 300, G914-926.

[120] Baumgartner W, et al. (2008) Heterotypic trans-interaction of LI- and E-cadherin and their localization in plasmalemmal microdomains. J. Mol. Biol. 378, 44-54.

[121] Bartolmas T, Hirschfeld-Ihlow C, Jonas S, Schaefer M, Gessner R (2012) LI-cadherin cis dimerizes in the plasma membrane Ca(2+) independently and forms highly dynamic transcontacts. Cell Mol. Life Sci. 69, 3851-3862.

Table 1. Representative members of the cadherin superfamily in humans. Members of the *Homo sapiens* cadherin superfamily are grouped in families and subfamilies according to multiple criteria (following [13]). Families defined through EC1-based alignments are also shown [20], with classification outliers highlighted in red.

Superfamily	Family	Subfamily	Name	Repeats	Peculiarities
Cadherin	C-1	Type-I	CDH1 (E) CDH2 (N) CDH3 (P) CDH4 (R) CDH15 (M)	5	Pro-domain
		Type-II	CDH5 (VE) CDH6 (K) CDH7 CDH8 CDH9 (T1) CDH10 (T2) CDH11 (OB) CDH12 (N2) CDH18 CDH19 CDH20 CDH22 CDH24	5	Pro-domain
		Desmosomal	DSC1 DSC2 DSC3 DSG1 DSG2 DSG3 DSG4	5	Pro-domain
		7D-family	CDH16 (Ksp) CDH17 (Li)	7	
		Solitary	**CDH13 (T)** **CDH26**	5 5	Pro-domain / Lacks TM and cyto-domain Pro-domain
	C-2	Flamingo/CELSR	CELSR1-3	9	Pro-domain / 7 TM helices
Cadherin-related	Cr-1a	Clustered protocadherins	PCDHα1-13; **αc1,2** PCDHβ1-16 PCDHγa1-12, γb1-7, **γc3-5**	6	α and γ subfamily members share a constant domain each in the cytoplasmic region
			PCDH12 (VE2) **PCDH20**	6 7	
		δ1	PCDH1 (AXPC) PCDH7 (BH,NF) PCDH9 PCDH11X PCDH11Y	7 7 7 7 7	aa insertion
		δ2	**PCDH10 (OL)** PCDH17 PCDH18 PCDH19 PCDH8 (arcadlin, PAPC)	6 6 6 6 6	aa insertion aa insertion aa insertion
	Cr-1b	ε	RET **DCHS1** **FAT4**	4* 27 34	
	Cr-2		**CDHR1 (PCDH21)** CDHR2 (PCDH24) CDH23 **CDHR5 (μPCDH)**	6 9 27 4	
	Cr-3		CLSTN1-3 PCDH15 **CDHR3** (CDH28) **CDHR4** (CDH29) FAT1 FAT2 FAT3	2 11 6 6 34 34 34	
	Solitary		**DCHS2**	?	

Figure Legends

Figure 1. Architecture and structural diversity of cadherins. (a) Representative topology diagram of two cadherin repeats (Cdh23 EC1+2). A typical EC repeat features seven β-strands labeled A to G. Three calcium ions (sites 1-3) are located at the linker region between repeats (green spheres). The ribbon diagram on the right depicts the 3D structure of the two repeats. (b) Detail of calcium binding sites at linker region between repeats. Protein side chains and backbone atoms are in stick representation for amino acids indicated within a calcium binding motif. (c) Arrangement of EC repeats for different family members. Classical and desmosomal cadherins feature a cleavable prodomain, five EC repeats, and variable cytoplasmic domains that bind catenins. Clustered protocadherins sport six EC repeats and variable cytoplasmic domains. Non-clustered protocadherins feature from two up to 34 EC repeats, and some of them have variable non-cadherin extracellular domains. (d)&(e) Illustration of *cis* vs *trans* interactions for classical cadherins mediating contact between two cells. *Trans* interactions facilitate the formation of *cis* complexes [22].

Figure 2. Strand-swapped and X-dimer *trans* interactions of classical cadherins [22]. (a) Structure of entire extracellular domain (EC1-5) Cdh2 (N-cadherin) engaged in a strand-swapped dimer (EC1-to-EC1 contact). Inserts show details of the exchange of tryptophans at position 2 (W2). Availability of W2 for binding is modulated by calcium binding (green spheres). (b) Cdh13 (T-cadherin) EC1+2 structure shown for two protomers in a X-dimer conformation. One protomer is shown in opaque surface representation, while the other is in cartoon and transparent surface representations. Inset shows details of the interaction. (c) Classical cadherins engage in a similar interaction that is thought to lead to the strand-swapped dimer (right) and facilitate *cis* interactions [22].

Figure 3. Inner-ear cadherin handshake and heterotypic cadherin complexes. (a) Hair cell bundle, made of actin-filled stereocilia arranged in a staircase, moves upon mechanical stimulation thereby applying tension to tip links (black box). (b) Schematic of an heterotetrameric complex of Cdh23 and Pcdh15 forming the tip link [75]. Inset shows details of interaction, with the tips of both proteins engaged in a "handshake" interaction bond (right) [79]. (c) Ribbon diagram (left) and surface representation (right) of the handshake interaction, showing calcium ions (green spheres), two protrusions (arrow heads), and the side chain of Arginine 113. (d) Detail of handshake interaction. (e) Detail of Pcdh15 N-terminus with a disulfide bond. (f) Detail of N-terminal calcium binding site 0 in Cdh23. (g-i) Hypothetical models for heterotypic interactions: (g) clustered protocadherins heterotetramers [95], (h) classical and delta-cadherin complexes [15], and (i) Cdh17 and Cdh1 [120,121].

Figure I. Domain organization and cytoplasmic partners for protocadherins involved in signaling. (a) Clustered protocadherins are divided in three subfamilies: α, β, and γ. Distinct constant cytoplasmic regions are present for all members of the α and γ subfamilies. Cytoplasmic binding partners are listed. Number of members for each family in the human genome is indicated in parenthesis (b) The non-clustered δ protocadherins feature seven and six EC repeats for subfamilies δ1 and δ2, respectively. Cytoplasmic domains have distinct sequence motifs (CM1, CM2 and CM3 for δ1 and CM1 and CM2 for δ2, and WIRS for α and δ2). Binding partners are listed. (c) Pcdh15 and Cdh23 have unusually long extracellular domains (11 and 27 EC repeats respectively). The two proteins and their cytoplasmic binding partners are involved in inherited deafness and blindness (Usher syndrome).

Figure II. Sequence relationships among *Homo sapiens* cadherin superfamily members. (a) The sequences of all cytoplasmic domains of human cadherins aligned using MUSCLE. Tree generated using Clustal and TreeDyn. (b) Classification using sequences of repeats EC1 to EC3 aligned as in (a). Calsyntenins and RET excluded from these analyses. Non-clustered protocadherins are clearly segregated in subgroups in the EC1-EC3 based classification (Cr-2, Cr-3, δ1, δ2, and δ2'). (c) Cadherins with potential for homotypic and heterotypic *trans* interactions within a subtree are highlighted in the human EC1-3 tree. (d,e) Cadherins with potential for *trans* and *cis* heterotypic interactions across sibling subtrees are highlighted in (d) and (e), respectively. Members of the clustered protocadherins form *cis* heterotypic complexes within their subtrees as well.

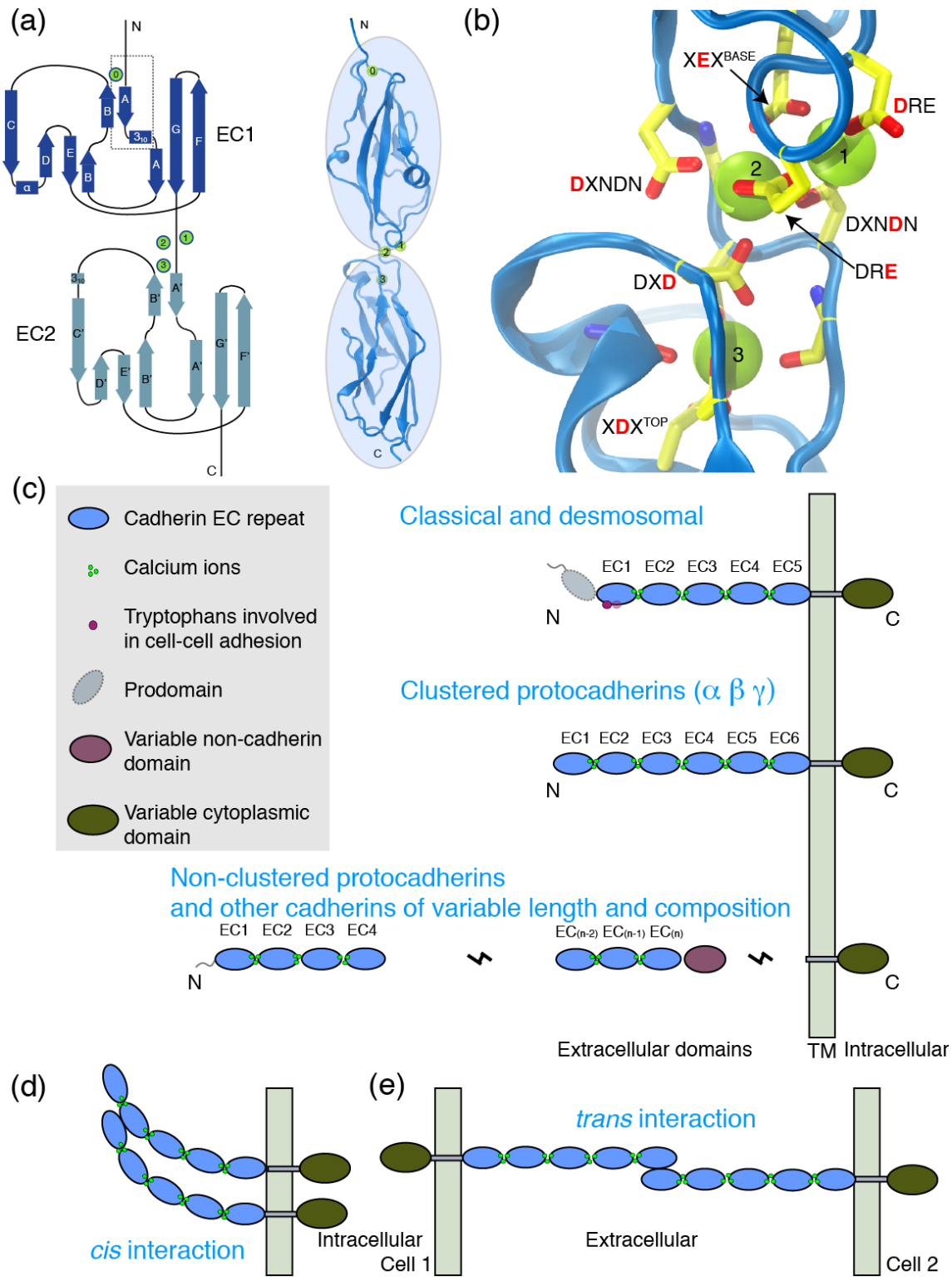

(a)
N
EC1
EC2
C
(b)
XEX BASE
DRE
DXNDN
DXNDN
DRE
DXD
XDX TOP
(c)
Cadherin EC repeat
Calcium ions
Tryptophans involved in cell-cell adhesion
Prodomain
Variable non-cadherin domain
Variable cytoplasmic domain
Classical and desmosomal
EC1 EC2 EC3 EC4 EC5
Clustered protocadherins (α β γ)
EC1 EC2 EC3 EC4 EC5 EC6
Non-clustered protocadherins
and other cadherins of variable length and composition
EC1 EC2 EC3 EC4
EC(n-2) EC(n-1) EC(n)
Extracellular domains
TM
Intracellular
(d)
cis interaction
Intracellular
Cell 1
(e)
trans interaction
Extracellular
Cell 2

Figure 1

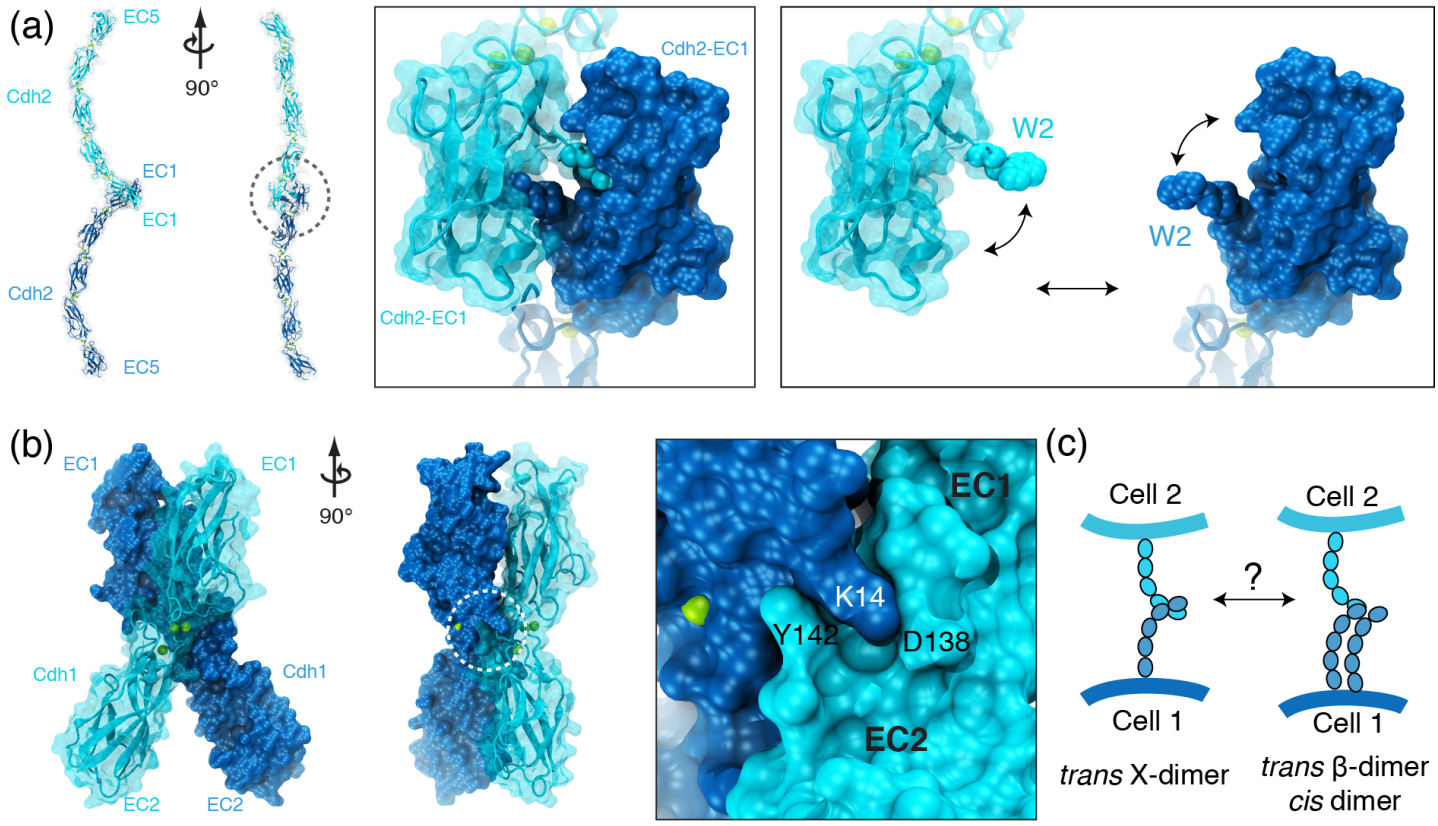

(a)
EC5
90°
Cdh2
EC1
EC1
Cdh2
EC5
Cdh2-EC1
Cdh2-EC1
W2
W2
(b)
EC1
EC1
90°
Cdh1
Cdh1
EC2
EC2
EC1
K14
Y142
D138'
EC2
(c)
Cell 2
Cell 2
?
Cell 1
Cell 1
trans X-dimer
trans β-dimer
cis dimer

Figure 2

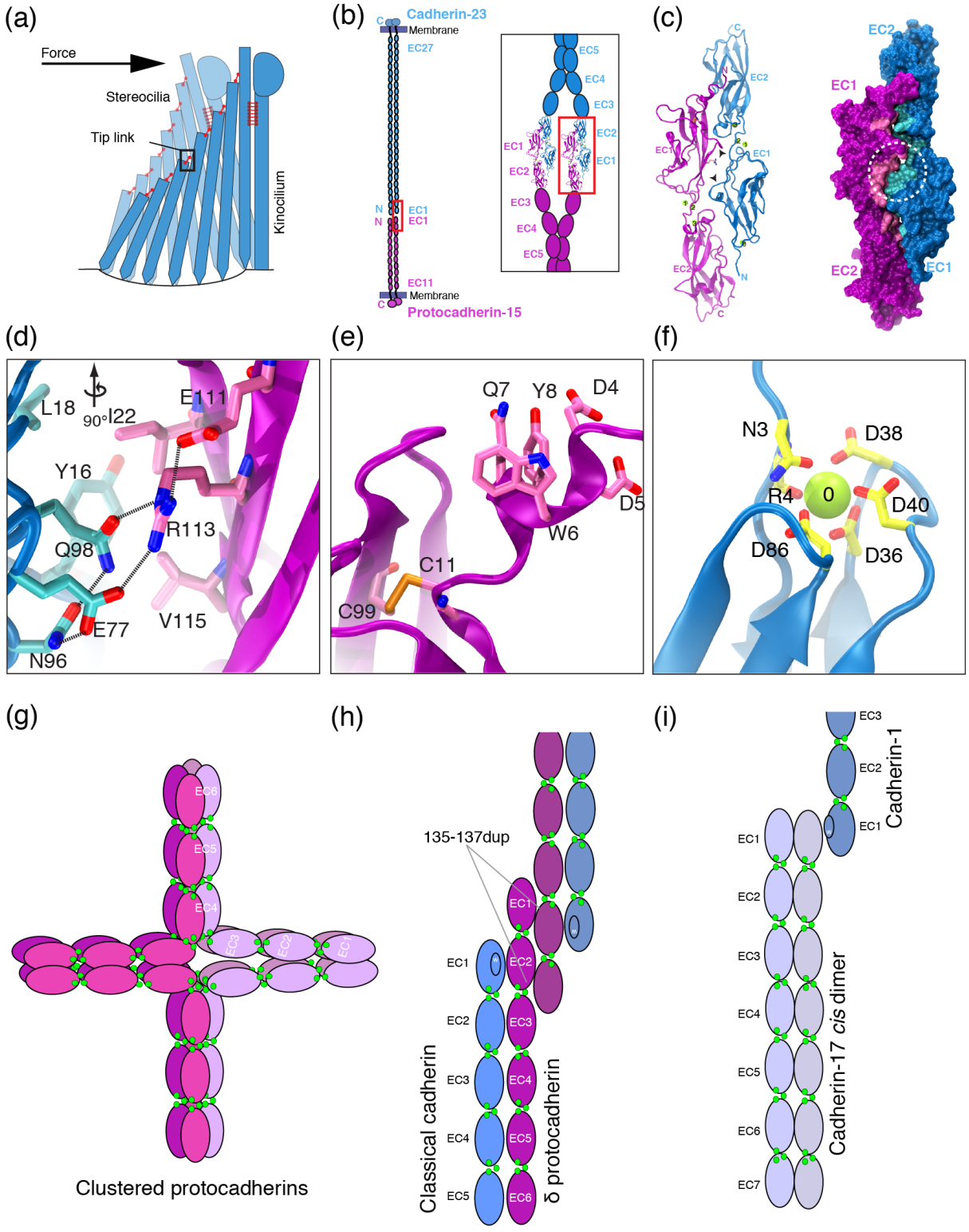

(a)
Force
Stereocilia
Tip link
Kinocilium
(b)
Cadherin-23
Membrane
EC27
EC1
EC11
Membrane
Protocadherin-15
EC5
EC4
EC3
EC2
EC1
(c)
EC2
EC1
(d)
L18
90°
I22
E111
Y16
R113
Q98
E77
V115
N96
(e)
Q7
Y8
D4
D5
W6
C11
C99
(f)
N3
D38
R4
0
D40
D86
D36
(g)
Clustered protocadherins
(h)
135-137dup
Classical cadherin
δ protocadherin
(i)
Cadherin-1
Cadherin-17 cis dimer

Figure 3

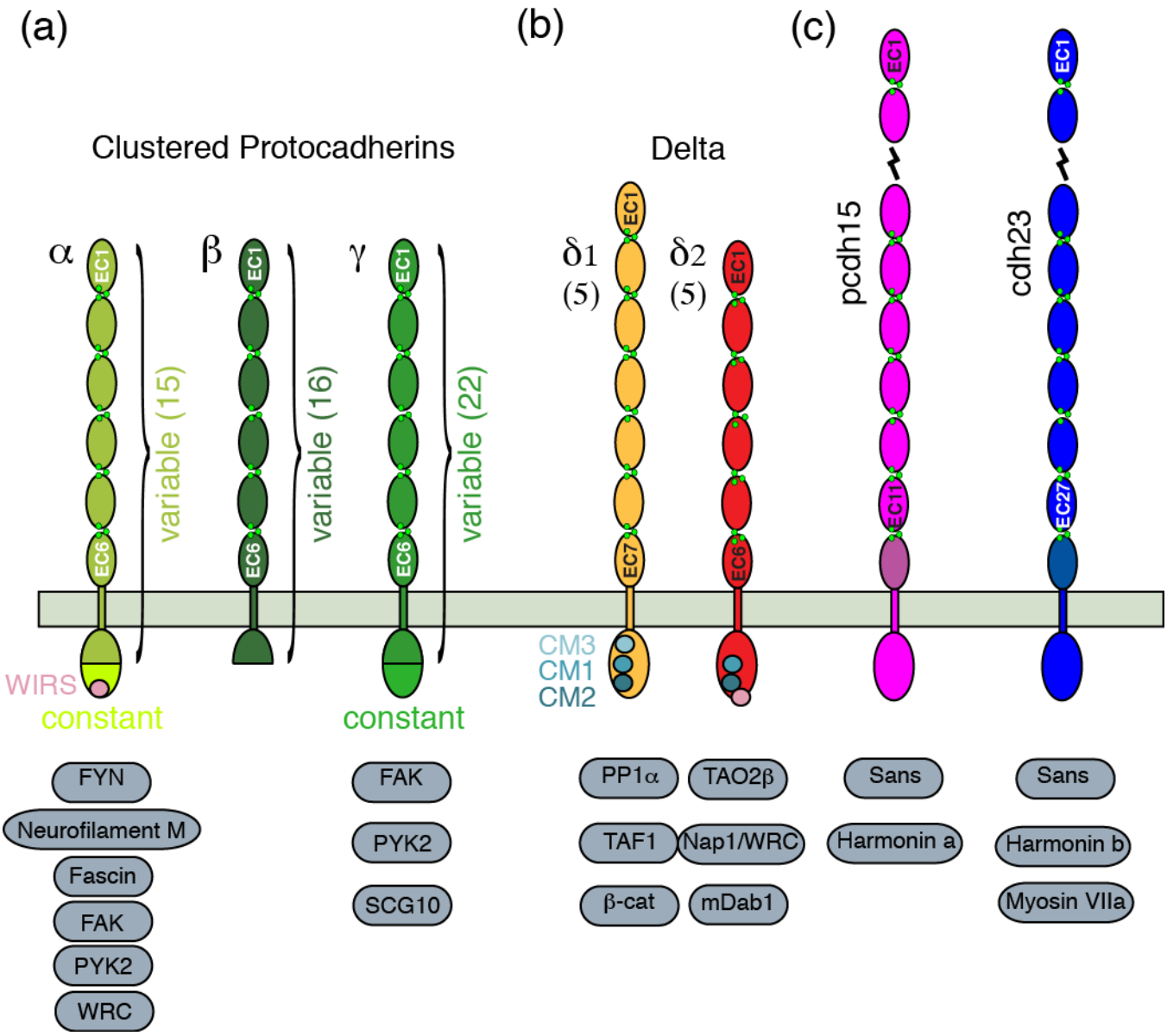
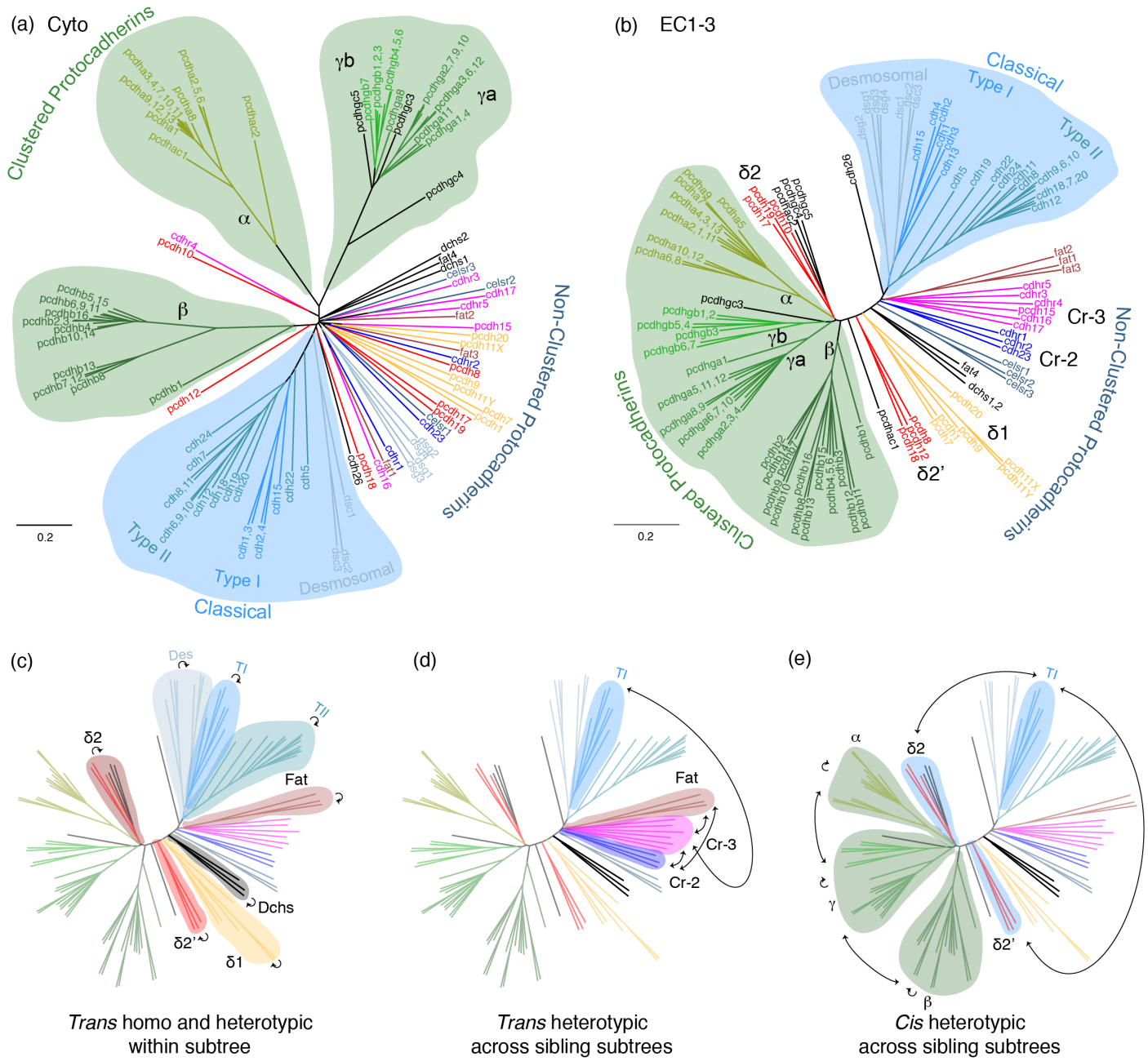

(a)
(b)
(c)
Clustered Protocadherins
Delta
α
β
γ
δ1
(5)
δ2
(5)
pcdh15
cdh23
EC1
EC6
EC7
EC11
EC27
variable (15)
variable (16)
variable (22)
WIRS
constant
constant
CM3
CM1
CM2
FYN
Neurofilament M
Fascin
FAK
PYK2
WRC
FAK
PYK2
SCG10
PP1α
TAO2β
TAF1
Nap1/WRC
β-cat
mDab1
Sans
Harmonin a
Sans
Harmonin b
Myosin VIIa

Figure I



Trans homo and heterotypic within subtree

Trans heterotypic across sibling subtrees

Cis heterotypic across sibling subtrees

Figure II